

**Action Research to Develop Effective
Physical Fitness Testing Standards for Pre-Service Physical Educators**

Abstract

Physical educators are often held to a higher standard of physical fitness. The ability to effectively convey the importance of physical fitness may depend upon the ability to appear physically fit. The ability to perform at a minimal level of proficiency on fitness tests was deemed important by the faculty of on physical education teacher education program (PETE). In an action research evaluation, the faculty examined standards presented in the literature as well as questionnaire responses by students to develop reasonable passing scores for physical education majors. Results indicated the students are receptive to minimal standards and the initial standards are achievable by most students in the PETE program.

Keywords: fitness testing, PETE program, fitness

Employing Action Research to Develop Effective Physical Fitness
Testing Standards for Future Physical Educators

The development of physical fitness is a goal for most K-12 physical education programs. Standards at state and national levels stress the importance of ensuring students achieve and maintain a health-enhancing level of physical fitness (Alabama Department of Education [ALSDE], 2012b; National Association for Sport and Physical Education [NASPE], 2013). Physically fit children are likely to enjoy an improved quality of life. These children are also less likely to be obese or to suffer the early stages of various metabolic disorders. Further, they tend to have more self-confidence, better mental health, and better academic performance (Janssen & LeBlanc, 2010; Strong et al., 2005). While some studies report moderate correlations between childhood and adult physical fitness, stronger correlations exist between a sedentary lifestyle during childhood and one during adulthood. Obese children are more likely to be obese adults, and fit children are more likely to be fit adults (Baranowski et al., 2000; Janz, Dawson, & Mahoney, 2000; Telama et al., 2005; Trudeau, Laurencelle, & Shephard, 2004). Given the value placed on the acquisition of fitness and fitness knowledge, assessment is inevitable. Welk and Meredith (2008) contended that fitness testing is important for a variety of reasons: 1) it stresses the importance of fitness for all, 2) it provides information teachers can use to make informed program changes, 3) it teaches children valuable self-assessment skills, 4) it aids in goal setting, and 5) it educates students and parents about the cognitive aspects of fitness and how it is developed. Some experts strongly question the value of top-down, mandated fitness testing in schools (Harris, 2006; Wrench & Garrett, 2008); however, criticisms of K-12 fitness testing tend to revolve around the inappropriate use of scores and a failure to link assessment with instruction. Fewer appear to be critical of the actual practice of assessing fitness. Whether or not

one agrees with the value or need for fitness testing, the reality is such that physical educators must be knowledgeable on the matter. In some states, state-wide physical fitness testing in public schools is required (Morrow Jr. & Ede, 2009). Most university physical education teacher education (PETE) majors (to be identified as *candidates*) in PETE programs will have the responsibility of helping K-12 students achieve and maintain health-enhancing levels of fitness. One could argue that these candidates need to be able to demonstrate their own maintenance of health-enhancing fitness levels.

Some PETE faculty have expressed the belief that candidates should undergo the same tests they will one day administer to students (Baghurst & Bryant, 2012). Through this process of *learning by doing*, candidates may gain a deeper understanding of fitness concepts and test procedures. Completing tests that are similar to those administered to students may allow candidates to better understand the emotions of student test-takers, as well as, perhaps, to identify test problems and pitfalls that would not otherwise be apparent. The repetition should also make the tests more familiar and provide candidates with the skills and confidence needed to reliably administer the tests to groups of students. Universally, all teachers, regardless of the content they teach, are well-served to “walk a mile” in their students’ testing shoes. When teachers in all areas become better acquainted with the various tests required, the entire teaching profession is strengthened.

Candidates should be conscious of their own fitness levels because they are often held to higher standards of fitness than other education professionals. Whitley, Sage, and Butcher (1988) reported that physical education teachers engaged in regular cardiovascular fitness were more likely to include aerobic conditioning content in their curricula. Researchers have shown that a physical education teacher’s body fat percentage can influence students’ willingness to accept

fitness information and engage in exercise (Melville & Maddalozzo, 1985). To measure this, researchers had the same teacher deliver fitness information via video to two randomly selected groups of students. The only difference between the groups was that, in one video, the teacher was wearing a fat suit. The students who watched the fat suit version of the video scored lower on retention tests and tended to view the teacher as less likable and as having less expertise. Spittle, Petering, Kremer and Spittle (2012) reported that pre-service physical education teacher self-perceptions of fitness and appearance are related to physical education teacher stereotypes. Their study showed that fit and physically competent physical educators tend to be viewed as more knowledgeable. Finally, physical educators who maintain a high degree of fitness are more likely to lead better qualities of life, to better perform daily rigorous work activities, and to better serve as effective and credible role models (Hetland & Strand, 2010; Staffo & Stier, 2000).

Role modeling is a valuable tool for motivating and informing learners. Unfortunately, just as students can be positively influenced by healthy model behavior, they can be negatively influenced by unhealthy model behavior. Administrators charged with hiring physical education teachers also believe in the importance of positive role modeling. Jenkins, Garn, and Jenkins (2005) reported that overweight job candidates with high GPAs were less favored by administrators than normal-weight job candidates with average GPAs. Many people incorrectly believe that fat always equals unfit (Blair, Cheng, & Holder, 2001; Blair & Church, 2004; O'Donovan et al., 2010). Physical inactivity does contribute to weight gain (Janssen & LeBlanc, 2010; LaMonte & Blair, 2012), and students, parents, and administrators may incorrectly assume that a fat physical education teacher is lazy and physically inactive.

PETE candidates will eventually be responsible for teaching fitness, which, in most schools, will involve the administration of fitness tests. Thus, an important function of physical

educators is to effectively administer fitness tests for program improvement and greater fitness awareness (Welk, 2008). Ideally, as candidates progress through a PETE program, they will grow familiar with the benefits of fitness testing and the challenges associated with fitness test administration. Some students have negative feelings towards fitness testing. Candidates should therefore be aware of potential student anxiety, embarrassment, and nervousness during the testing process (Wrench & Garrett, 2008). Going through this process themselves should help candidates develop a sense of empathy and learn to think critically about ways to alleviate the stress that accompanies fitness testing.

Action Research

We believed the most effective form of study for our research setting to be an action research methodology, which was defined by Parsons and Brown (2002) as follows:

Action research is a form of investigation designed for use by teachers to attempt to solve problems and improve professional practices in their classrooms. It involves systematic observations and data collection which can be used by the practitioner-researcher in reflection, decision-making and the development of more effective classroom strategies.
(p. 55)

With the goal of identifying fitness test cut-off scores that are effective and fair for candidates, we sought to promote change. The inclusion of stakeholders was important because the decisions to be made were intended to change candidates' habits and attitudes towards fitness. Brydon-Miller, Greenwood, and Maguire (2003) suggested that one goal of action research is to generate knowledge that will help advance community and individual well-being.

Action research has been described as a process whereby communities reflect upon and improve their own practice by linking reflection and action. For this study, the action research

model followed was described by Altrichter, Kemmis, McTaggart, and Zuber-Skerritt (2002) as consisting of four phases: 1) planning, 2) acting, 3) observing, and 4) reflecting.

Given the assumption that a fitness testing requirement has the potential to better prepare physical educators, this action research study was conducted to answer two questions: 1) *What are candidates' current attitudes and beliefs regarding fitness testing?* and 2) *What are reasonable cut-off scores with which to begin in our particular setting?* In an effort to answer these two questions, we initiated a review of fitness test criteria, administered the current state fitness test for grades 3 through 12 to our candidates and analyzed the results, and gathered candidate input regarding attitudes and preferences concerning fitness testing as a PETE program requirement.

Planning

Review of Published Fitness Values

The planning phase began with a review of previously published fitness standards. Normative standards and cut-off scores used in other situations were reviewed to establish initial valid cut-off scores for our particular setting. Fitness test values and coinciding evaluative categories were obtained for: 1) the Fitnessgram (Mahar & Rowe, 2008; Plowman & Yan-Shu Liu, 1999; Welk, Going, Morrow, & Meredith, 2011), 2) the American College of Sports Medicine (ACSM) Fitness Test (Dwyer & Davis, 2008), and 3) a fitness test created by the Canadian Society for Exercise Physiology (2003). The PETE faculty reviewed both the published values and the candidates' scores from prior semesters and agreed upon suggested pass/fail cut-off scores to propose to the candidates. When comparing the fitness subtest scores needed for an individual to be categorized as *healthy* (normally the 20th to 25th percentile), we elected to go with the lowest value. For example, when comparing the number of laps needed on

the PACER test for females to be considered *healthy*, one test reported 33 laps, another 36 laps, and the third 41 laps. We thus used 31 laps for our cut-off value, assuming that, in the future, it would be preferable to raise standards than to lower standards.

Creating a Questionnaire for Our Students

In addition to demographic information, the questionnaire attempted to gather data on PETE students' attitudes and perceptions of mandatory fitness testing. The questionnaire contained three parts: two topics on a bipolar adjective scale (eight adjective pairs), six rating scale questions, and three open-ended questions.

The two bipolar adjective topics posed to the candidates were: 1) *physical fitness tests* and 2) *the practice of making physical education majors pass a physical fitness test as a graduation requirement*. The participants were asked to think about the topic statement and respond by marking the scale to show how they felt. The seven point scales ranged as follows: fair-unfair, harmful-helpful, necessary-unnecessary, negative-positive, useful-useless, discouraging-encouraging, valuable-worthless, bad-good. The eight adjective pairs were instrumental (e.g. useful-useless; helpful-harmful), rather than affective (e.g. exciting-boring; pleasant-unpleasant) in nature because such adjective types have been successfully used in the physical semantic differentials that assess attitudes towards physical activity (Rhodes, Hunt Matheson, & Mark, 2010).

The second part of the questionnaire contained questions requiring responses on a rating scale that ranged from *strongly disagree* to *strongly agree*. Questions asked about current fitness levels, opinions of physical education teachers as role models, and perceptions of the locus of control with respect to fitness.

The final section contained three open-ended questions: 1) In your opinion, what is the longest amount of time it should take a 20-year-old physical education major to run a mile? 2) Assume this university instituted a mandatory fitness test for all PETE majors. What should happen to students who fail? 3) Make any additional comments you would like on this topic.

Action

The action phase consisted of administering the fitness test and the questionnaires to the PETE candidates. In both instances, the research questions were shared with the candidates because they were participants in the study and, as participants, they were allowed to see the results and provide feedback at any point. The goals of promoting positive change in our teacher preparation program were explicitly stated, and efforts were made to create a collaborative environment.

Candidates

Data were gathered following Institutional Review Board approval. Participation in the fitness test was part of a required class activity. The submission of scores and completion of the questionnaire were voluntary. All participants signed a consent form.

The participating candidates consisted of physical education majors at a mid-sized regional university in the southeastern United States. Enrollment is approximately 9,000 students, with 85% of the students enrolled in undergraduate programs. Males comprise 40% and females comprise 60% of the student population. African Americans comprise 30%, white 65%, and all other ethnicities less than 5%. The majority of the PETE students are considered traditional college age (between 18 and 22 years old). Participants in the study ranged from 18 to 39 years, but the vast majority were under 24 years of age.

Fitness Testing of Candidates

The fitness test was administered to the candidates as part of a class exercise. The four test components administered were the V-sit, the progressive aerobic cardiovascular endurance run (PACER), the push-ups and the curl-ups. The procedures described in the state test manual were followed (ALSDE, 2012a), and candidates were given an explanation of the test and a demonstration of how to perform each test component. Candidates were also shown the proposed cut-off scores created by the PETE faculty (values shown in table 1). Though the candidates were informed that these cut-off scores were not yet in effect, they were encouraged to attempt to surpass the cut-off scores as a means of self-assessment. Due to absenteeism, injury and non-compliance, not every candidate completed every test component.

Table 1

Departmentally proposed cut-off values for PETE candidates.

Age	PACER (laps)		Push-ups		Curl-ups		V-sit (inches)	
	<i>Men</i>	<i>Women</i>	<i>Men</i>	<i>Women</i>	<i>Men</i>	<i>Women</i>	<i>Men</i>	<i>Women</i>
29 & under	51	31	14	6	17	16	13	16
30 to 39	45	29	13	6	15	15	13	15
40 to 49	32	24	12	5	14	14	12	12
50 & over	22	19	8	3	12	9	10	8

Candidate Questionnaires

The final step in the data gathering process was to gather information on candidate attitudes and opinions regarding fitness testing and, more specifically, the use of mandatory cut-off passing values. Because some candidates could not participate in the fitness test due to injury or illness, the N value for the questionnaire ($N = 48$) was greater than that for the fitness test ($N =$

43). Questionnaires were administered to three classes by department faculty. The average age was 22.4 ± 5.3 , and 44% of the candidates were female.

Data Analysis

While the establishment of cut-off values was a subjective exercise, the values were not chosen arbitrarily. Instead, the faculty chose the values based on a review of the candidate scores and the normative values associated with the bottom quintile of values reported for the Fitnessgram, the ACSM Fitness Test, and the Canadian Fitness Test. The candidate scores were then compared to the proposed cut-off values to see what the passing rate would have been had the cut-off values been in effect.

The questionnaire data were entered into an SPSS version 21 (SPSS IBM, New York, U.S.A) spreadsheet and reverse coded where necessary so that higher scores on a 1 to 7 or 1 to 5 scale were always indicative of a more desirable belief or attitude. The descriptive data were computed, and a principle component analysis (CPA) was performed on the semantic differential topics to examine their usefulness and to identify potential areas for data reduction. The CPA explored two attitude topics: 1) *physical fitness tests* (Test Attitude) and 2) *the practice of making physical education majors pass a physical fitness test as a graduation requirement* (Mandatory Test Attitude). The rating scale questions were totaled and overall means were compared according to gender and perceived current fitness level. Lastly, the three open-ended questions were organized: responses about an appropriate mile run cut-off were simply summed, responses regarding the suggested procedure to follow in the event of test failure were grouped by theme, and responses for any additional comments were coded as positive, negative, or neutral.

Observing

Fitness Test Results

A total of 43 students, 16 (37%) female and 27 (63%) male, completed the test during a required physical education course. The mean age of the participants was 22.8 ± 5.8 . Prior to test administration, candidates were shown the newly proposed cut-off scores and advised that a passing score would eventually be needed to graduate. Passing rates for each test component were then calculated using the cut-off scores; these are shown in table 2.

Table 2

Passing rates with proposed HPER department fitness test cut-off scores.

	PACER	Push-ups	Curl-ups	V-sit
Male ($n = 27$)	71%	86%	93%	100%
Female ($n = 16$)	43%	100%	100%	100%

Questionnaire Results

The responses to the bipolar adjective topics for both concepts, Test Attitude and Mandatory Test Attitude, were positive. For both topics, the responses were generally consistent when compared with the various adjective pairs. For Test Attitude, the only adjective pair with lower scores than the rest was the *encouraging–discouraging* pair. For Mandatory Test Attitude, all eight adjective pairs were similar. The semantic differentials for both the 8-item Fitness Test object ($\alpha = .871$) and the 8-item Mandatory Fitness Test object ($\alpha = .944$) were found to be highly reliable. Descriptive statistics, presented in table 3, indicated that respondents held positive attitudes toward both fitness tests in general and the practice of mandating fitness tests for physical education majors in particular. A visual inspection of the means indicated

overwhelmingly positive attitudes and confirmed the t-tests. A comparison of means indicated that both fitness tests ($t(47) = 9.24, p < .001$) and mandatory fitness test ($t(47) = 7.18, p < .001$) totals were significantly greater than a neutral score.

Table 3

Descriptive statistics for bipolar adjective pairs.

Semantic differential items	Mean	SD
Physical fitness tests		
Fitness tests are fair/unfair	5.69	1.82
Fitness tests are helpful/harmful	5.77	1.51
Fitness tests are necessary/unnecessary	5.69	1.61
Fitness tests are positive/negative	5.52	1.64
Fitness tests are useful/useless	5.88	1.77
Fitness tests are encouraging/discouraging	4.96	1.52
Fitness tests are good/bad	5.65	1.44
Fitness tests are valuable/worthless	5.60	1.87
Overall	5.59	1.65
The practice of making physical education majors pass a physical fitness test as a graduation requirement		
Fitness test requirements are fair/unfair	5.77	2.04
Fitness test requirements are helpful/harmful	5.75	1.72
Fitness test requirements are necessary/unnecessary	5.38	1.96
Fitness test requirements are positive/negative	5.56	1.75
Fitness test requirements are useful/useless	5.58	1.90
Fitness test requirements are encouraging/discouraging	5.46	1.76
Fitness test requirements are good/bad	5.71	1.71
Fitness test requirements are valuable/worthless	5.73	1.83
Overall	5.62	1.84

Note: N = 48. Responses were given on a 7-point scale, with 1 being the most negative response and 7 being the most positive response.

The descriptive data for the rating scale questions revealed generally positive responses regarding fitness testing for physical education majors. Using a 5-point scale ranging from

strongly disagree to strongly agree, with higher values indicating greater favorability, the median and mean values for all six questions were more positive than a neutral score of 3.0. The descriptive data are shown in table 4.

Table 4

Descriptive statistics for Likert-type rating scale questions.

Likert Type Question	N	Median	Mean	SD
It is important that physical education teachers be able to demonstrate a high level of fitness.	48	5	4.21	1.15
A minimum level of physical fitness should be required of physical education teachers.	48	4	3.48	1.35
For a physical education teacher, being physically fit makes one a better role model.	48	5	4.56	0.9
Knowing a physical fitness test must be passed in order to graduate will motivate me to engage in regular health-enhancing physical activities.	48	5	4.33	1.1
^a Physical education majors have little control over their own physical fitness levels.	48	4.5	4.23	1.02
My current physical fitness level is satisfactory for a physical educator.	47	4	3.89	1.37

Note: Responses were given on a 5-point scale, with 1 being *strongly disagree* and 5 being *strongly agree*. ^aSince, on this item, a low score indicates a positive perception, the scores for this item were reversed when the data were analyzed (1=strongly agree and 5=strongly disagree).

A one-way ANOVA showed that the perceptions of males and females did not differ. Similarly, perceptions of fitness testing were not influenced by the respondents' current (albeit self-identified) fitness levels. Lastly, the total scores for all rating scale questions were significantly higher than what would have been observed if the respondents had selected *neutral* for each item. In other words, opinions were clearly positive.

The final three questions on the questionnaire were open-ended. The responses to these questions are represented graphically in figure 1. When asked about the highest acceptable mile-run time limit for a physical education major, participants replied with a median time of roughly

10 minutes. When asked what should happen to candidates who failed the fitness test, most candidates favored giving them more chances to pass.

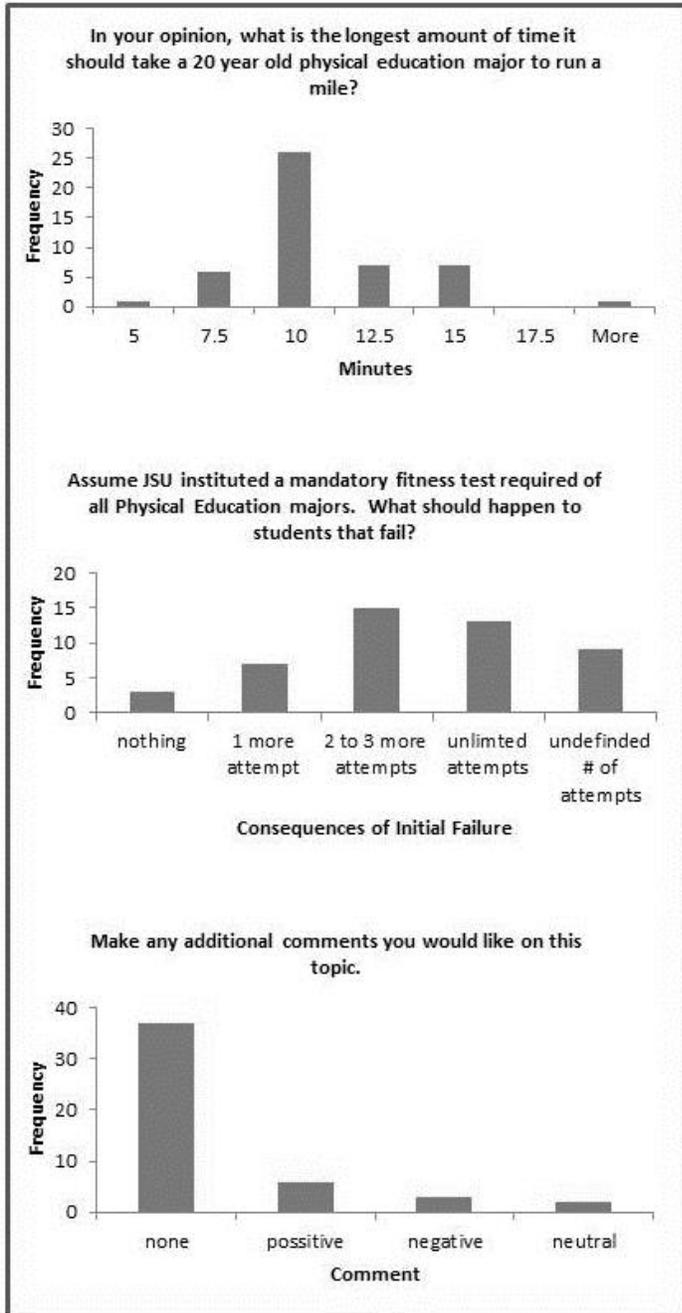


Figure 1. Responses to open-ended questions

Nearly 88% of the respondents believed that candidates should have to pass a physical fitness test; however, they also believed that candidates should be allowed multiple attempts, with suggestions ranging from one additional attempt to unlimited attempts. When given a chance to opine about any matter relating to the study, only 11 respondents (23%) chose to do so. Six responses were classified as positive, three as negative, and two as neutral. Negative comments indicated that fitness testing was a “waste of time” or “unfair.” Positive comments generally supported fitness testing as a means of raising standards and promoting professionalism. The two neutral comments offered suggestions regarding procedural improvements for test administration.

Reflecting

The results indicated that physical education majors in the studied PETE program appeared supportive of efforts to implement a fitness testing component. Evidence further indicated that candidates were willing to make the testing requirement a positive accountability measure. An examination of the candidates’ actual fitness test scores revealed that the majority had the capacity to meet the minimum suggested standards. The component raising the greatest concern was the PACER for females, with a pass rate of only 43%. However, females performed extremely well on the other three components. Males performed extremely well on the v-sit and curl-ups, with room for improvement on the push-ups and PACER. As an indication that candidates can expend greater effort when challenged, once cut-off scores were shared prior to testing, the male PACER and v-sit scores improved significantly from previous semesters. Thus, with more incentive to reach higher benchmarks, candidates may perform better. It is important to note that some candidates likely stopped performing once the cut-off scores were exceeded, even though they had not reached exhaustion and could have continued. In the future, it is

possible that a fitness certificate could be awarded to candidates who go beyond departmental standards. When the cut-off scores were shown to the candidates, their responses seemed to suggest that they want to be held accountable for maintaining a minimum level of fitness. This result was supported by the questionnaire responses. When asked about an appropriate minimum mile run time for physical education majors, the median suggestion was 10 minutes, which is equivalent to the minimum female PACER score of 30 laps (31 laps was the cut-off score for females under 30 yrs.). The display of positive attitudes and increased effort associated with the reveal of departmental cut-off scores may suggest that standards can be slowly increased over a few years as candidates acclimate to the new requirements and continue to support the inclusion of mandatory fitness standards.

The questionnaire revealed positive attitudes toward fitness tests and the possibility of mandating such tests for physical education majors. Both the rating scale questions and the semantic differential items indicated that students believe physical education teachers should maintain a minimum level of fitness to be effective professionals and competent role models. These views remained strong regardless of gender or perceived current fitness levels. One might expect students perceiving themselves as *less fit* to be less likely to have positive attitudes toward mandatory fitness tests, but this was not the case. It is hoped that with the introduction of mandatory fitness testing, even the less fit students, realizing the importance of maintaining their fitness, will be motivated to improve. We were encouraged to observe no difference in attitudes according to gender. Fitness tests in other arenas, such as the military and fire-fighting, have at times been viewed negatively as veiled attempts to force females to fail. The females in this study did not appear to have negative views of fitness testing and, in fact, had higher pass rates than the males.

A dialogue has already begun about directions for future practice. Given the cyclical or spiraling nature of action research (Altrichter et al., 2002), as researchers, we felt it necessary to consider what changes will be necessary going forward. Additionally, in action research, the sharing of results is critical. When discussing the cut-off fitness requirements, the faculty shared the study findings with candidates, who should be aware that their voices are heard regarding protocols and procedures. Candidates in this study supported multiple pass attempts for the fitness test; thus, all involved should work together to establish protocols to allow for formative assessment.

As stakeholders, PETE candidates should know that their input is valued and necessary throughout the action research process. The evidence in the action research showed that the PETE candidates believed fitness was important and that they believed it was part of being a professional in their chosen field. When introducing the mandatory fitness testing requirement to candidates, it will be imperative to stress the positive effect it will have on raising their levels of professionalism and their credibility in the K-12 community. Candidates should know that a minimum required level of fitness is not a punitive action being forced on them from above, but, rather, an attempt to raise standards that other candidates clearly support.

References

- Alabama Department of Education (ALSDE). (2012a). Alabama Physical Fitness Assessment: Test Administrator Manual. Retrieved from <http://www.alsde.edu/home/general/hpeinformation.aspx>
- Alabama Department of Education (ALSDE). (2012b). Courses of Study for Physical Education. Retrieved November 27, 2013, from <http://alex.state.al.us/standardAll.php?subject=PE1&summary=1>
- Altrichter, H., Kemmis, S., McTaggart, R., & Zuber-Skerritt, O. (2002). The concept of action research. *The Learning Organization*, 9(3), 125–131. doi:10.1108/09696470210428840
- Baghurst, T., & Bryant, L. C. (2012). Do as I Say, Not as I Do: Improving the image of the physical education profession. *Strategies*, 25(4), 11–13.
- Baranowski, T., Mendlein, J., Resnicow, K., Frank, E., Cullen, K. W., & Baranowski, J. (2000). Physical activity and nutrition in children and youth: an overview of obesity prevention. *Preventive Medicine*, 31(2), S1–S10.
- Blair, S. N., Cheng, Y., & Holder, J. S. (2001). Is physical activity or physical fitness more important in defining health benefits? *Medicine and Science in Sports and Exercise*, 33(6; SUPP), S379–S399.
- Blair, S. N., & Church, T. S. (2004). The fitness, obesity, and health equation. *JAMA: The Journal of the American Medical Association*, 292(10), 1232–1234.
- Brydon-Miller, M., Greenwood, D., & Maguire, P. (2003). Why Action Research? *Action Research*, 1(1), 9–28. doi:10.1177/14767503030011002

Canadian Society for Exercise Physiology (CSEP). (2003). *The Canadian Physical Activity, Fitness and Lifestyle Approach (CPAFLA) 3rd edition*. Canadian Society for Exercise Physiology.

Dwyer, G. B., & Davis, S. E. (Eds.). (2008). *ACSM's Health-Related Physical Fitness Assessment Manual* (2nd ed.). Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins.

Harris, J. (2006). A review of children's fitness testing. *European Physical Education Review, 12*(2), 201–225. doi:10.1177/1356336X06065359

Hetland, K. M., & Strand, B. (2010). A Descriptive Analysis of Undergraduate PETE Programs in the Central District. *The ICHPER-SD Journal of Research in Health, Physical Education, Recreation, Sport & Dance, 5*(1), 3–9.

Janssen, I., & LeBlanc, A. G. (2010). Review Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *International Journal of Behavioral Nutrition and Physical Activity, 7*(40), 1–16.

Janz, K. F., Dawson, J. D., & Mahoney, L. T. (2000). Tracking physical fitness and physical activity from childhood to adolescence: the Muscatine study. *Medicine and Science in Sports and Exercise, 32*(7), 1250–1257.

Jenkins, J. M., Garn, A., & Jenkins, P. (2005). Preservice teacher observations in peer coaching. *Journal of Teaching in Physical Education, 24*(1), 2-23.

LaMonte, M. J., & Blair, S. N. (2012). Physical Activity, Fitness, and Mortality Rates. In *Physical Activity and Health* (2nd ed., pp. 167–184). Champaign, IL: Human Kinetics.

Mahar, M. T., & Rowe, D. A. (2008). Practical Guidelines for Valid and Reliable Youth Fitness Testing. *Measurement in Physical Education and Exercise Science, 12*(3), 126–145.

doi:10.1080/10913670802216106

Melville, D. S., & Maddalozzo, J. G. (1985). The Effects a Physical Educator's Appearance of Body Fatness Has on Communicating Exercise Concepts to High School Students.

Morrow Jr, J. R., & Ede, A. (2009). Research Quarterly for Exercise and Sport Lecture Statewide Physical Fitness Testing: A BIG Waist or a BIG Waste? *Research Quarterly for Exercise and Sport*, 80(4), 696–701.

National Association for Sport and Physical Education (NASPE). (2013). NASPE National Standards for Physical Education. *Standards and Position Statements*. Retrieved November 27, 2013, from <http://www.aahperd.org/naspe/standards/nationalstandards/pestandards.cfm>

O'Donovan, G., Blazevich, A. J., Boreham, C., Cooper, A. R., Crank, H., Ekelund, U., ... Stamatakis, E. (2010). The ABC of Physical Activity for Health: a consensus statement from the British Association of Sport and Exercise Sciences. *Journal of Sports Sciences*, 28(6), 573–591. doi:10.1080/02640411003671212

Parsons, R., & Brown, K. (2002). *Teacher as reflective practitioner and action researcher*. Belmont, CA: Wadsworth/ Thomas.

Plowman, S. A., & Yan-Shu Liu, N. (1999). Norm-Referenced and Criterion-Referenced Validity of the One-Mile Run and PACER in College Age... *Measurement in Physical Education & Exercise Science*, 3(2), 63.

Rhodes, R., Hunt Matheson, D., & Mark, R. (2010). Evaluation of Social Cognitive Scaling Response Options in the Physical Activity Domain. *Measurement in Physical Education and Exercise Science*, 14(3), 137–150. doi:10.1080/1091367X.2010.495539

Spittle, M., Petering, F., Kremer, P., & Spittle, S. (2012). Stereotypes and self-perceptions of physical education pre-service teachers. *Australian Journal of Teacher Education*, 37(1), 2.

Staffo, D. F., & Stier, W. F. (2000). The Use of Fitness Tests in PETE Programs. *Journal of Physical Education, Recreation & Dance*, 71(5), 48–52.

doi:10.1080/07303084.2000.10605146

Strong, W. B., Malina, R. M., Blimkie, C. J., Daniels, S. R., Dishman, R. K., Gutin, B., ... Pivarnik, J. M. (2005). Evidence based physical activity for school-age youth. *The Journal of Pediatrics*, 146(6), 732–737.

Telama, R., Yang, X., Viikari, J., Välimäki, I., Wanne, O., & Raitakari, O. (2005). Physical activity from childhood to adulthood: a 21-year tracking study. *American Journal of Preventive Medicine*, 28(3), 267–273.

Trudeau, F., Laurencelle, L., & Shephard, R. J. (2004). Tracking of physical activity from childhood to adulthood. *Medicine and Science in Sports and Exercise*, 36(11), 1937–1943.

Welk, G. J. (2008). The Role of Physical Activity Assessments for School-Based Physical Activity Promotion. *Measurement in Physical Education and Exercise Science*, 12(3), 184–206.

doi:10.1080/10913670802216130

Welk, G. J., Going, S. B., Morrow, J. R., & Meredith, M. D. (2011). Development of New Criterion-Referenced Fitness Standards in the FITNESSGRAM® Program. *American Journal of Preventive Medicine*, 41(4), S63–S67. doi:10.1016/j.amepre.2011.07.012

Welk, G. J., & Meredith, M. D. (Eds.). (2008). *Fitnessgram / Activitygram Reference Guide*. Dallas, TX: The Cooper Institute.

Whitley, J. D., Sage, J. N., & Butcher, M. (1988). Cardiorespiratory Fitness: Role Modeling by P.E. Instructors. *Journal of Physical Education, Recreation & Dance*, 59(7), 81.

Wrench, A., & Garrett, R. (2008). Pleasure and pain: Experiences of fitness testing. *European Physical Education Review, 14*(3), 325–346. doi:10.1177/1356336X08095669

Zhu, W., Mahar, M. T., Welk, G. J., Going, S. B., & Cureton, K. J. (2011). Approaches for development of criterion-referenced standards in health-related youth fitness tests. *American Journal of Preventive Medicine, 41*(4), S68–S76.

Zhu, W., Plowman, S. A., & Park, Y. (2010). A Primer-Test Centered Equating Method for Setting Cut-off Scores - ProQuest. *Research Quarterly for Exercise and Sport, 81*(4), 400–409.